DOCUMENT RESUME

ED 258 935 SP 026 123

AUTHOR Fielding, Glen; Schalock, Del
TITLE The Development of a Teacher's Handbook and a Related Staff Development Program for Integrating Teaching and Testing in High School. Final Report.
INSTITUTION Oregon State System of Higher Education, Monmouth. Teaching Research Div.
SPONS AGENCY National Inst. of Education (ED), Washington, DC.
PUB DATE Jan 85
CONTRACT 400-82-0013
NOTE 77p.; For the resource guide, see SP 026 122; for the handbook, see ED 257 821.
PUB TYPE Reports - Descriptive (141)

EDRS PRICE MF01/PC04 Plus Postage.
DESCRIPTORS *Faculty Handbooks; Inservice Teacher Education; *Instructional Design; *Material Development; Secondary Education; *Secondary School Teachers; *Testing; *Test Use

ABSTRACT

This project had three major goals: (1) To develop a handbook for high school teachers that shows how instruction and testing can work together to foster student learning; (2) To pilot test this handbook in the context of a staff development program for high school teachers; and (3) To prepare training manuals that would allow the staff development program to be implemented in school districts that had not participated in its development or pilot testing. The handbook sets forth a goal-based approach to integrating teaching and testing which intends to reflect the contextual factors that influence teacher practices, e.g., student characteristics. The first section of this report presents an overview of the processes followed in developing the handbook, the issues faced in determining its design, and the distinctive characteristics of the final product. The second section describes the development and pilot-testing of a staff development program based on the handbook. In the final section, descriptions are offered of the content and organization of three manuals developed by the project: the "Resource Guide for Teachers," the "Planning Guide for Lead Teachers and Principals," and the "Program Guide for Superintendents." This section also contains a description of the refinement process and field trials that are planned for the manuals. Included in the appendixes are a review of the literature on integrating teaching and testing, and the questionnaire used to gather information about high school teachers' testing procedures. (JD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

SCOPE OF INTEREST NOTICE
The Eric Facility has assigned this document for processing to:

In our judgment, this document is also of interest to the Clearinghouses noted to the right. Indexing should reflect their special points of view

ED258935

THE DEVELOPMENT OF A TEACHER'S HANDBOOK AND A RELATED STAFF DEVELOPMENT PROGRAM FOR INTEGRATING TEACHING AND TESTING IN HIGH SCHOOL

FINAL REPORT

Contract # 400-82-0013
The National Institute of Education

Prepared by

Glen Fielding
Del Schalock

January 1985

The Teaching Research Division
Oregon State System of Higher Education

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
✓ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

SP026123

DC13

CONTENTS

	Page
Goals of the Project	1
Development of the Handbook	4
. Overview	4
. Issues faced in designing and drafting the Handbook	4
. Distinctive characteristics of the final product	10
Development and Pilot-Testing of a Staff Development Program Based on the Handbook	12
. Purpose	12
. Overview	13
. Expected Outcomes	14
. Training procedures	15
. Reactions of participants	17
. Reactions of lead teachers	22
. Use of the pilot test as a basis for designing research-related measures	23
Preparation of Staff Development Manuals for Teachers, Lead Teachers and Principals, and Superintendents	25
. The Resource Guide for Teachers	26
. The Planning Guide for Lead Teachers and Principals	26
. The Program Guide for Superintendents	28
. Plans for field testing and refining the staff development manuals	29
Relation to Current Work Suported by the NIE	30
References	32

Appendices

1. Members of the advisory panels that guided the development of the Handbook

2. A review of the literature on integrating teaching and testing.

3. A questionnaire used to gather information about high school teachers' testing practices.

GOALS OF THE PROJECT

This project had three major goals:

(1) To develop a handbook for high school teachers that shows how instruction and testing can work together to foster student learning;

(2) To pilot test this handbook in the context of a staff development program for high school teachers; and

(3) To prepare training manuals that would allow the staff development program to be implemented in school districts that had not participated in its development or pilot testing.

The development of a handbook on linking instruction and assessment seemed of crucial importance both because of the content with which it was to deal and the form it was to take. From the standpoint of content, the idea of a handbook appeared timely in view of mounting evidence that the quality of teaching and testing in high school courses is far from satisfactory. Studies of high school classrooms and teachers suggest the following conclusions.

1. Instruction across all subject areas generally is geared to the lowest levels of learning and consists largely of lectures, followed by question-answer sessions. Textbooks typically define the scope and sequence of instruction. Classroom assessment by and large consists of quizzes and objective tests that call for factual responses (Sirotnik, 1983).

2. Teachers on the whole are poorly trained in the area of assessment, and high school teachers receive even less training than elementary school teachers (Stiggins & Bridgeford, 1982). To the extent that staff development programs in this area are available, they generally focus on administering and interpreting tests mandated and managed by either the school district or the state (Burry et al., 1982).

3. Without specific training, teachers are apt to have difficulty in designing assessment tools that are tied closely to instructional goals. Indeed, a recent study suggests that secondary school teachers may have difficulty in determining whether a given set of test items corresponds or fails to correspond to a given learning goal (Carter, 1984).

4. It appears that classroom tests are used primarily for purposes of grading. Teachers do not regularly use tests as a guide to instructional planning and decision-making or as an aid to student learning (Rudman et al., 1980) Students receive much less feedback on their learning progress than the literature on effective instruction in general and mastery learning in particular suggests is desirable (Sirotnik, 1983).

The proposed content of the handbook thus appeared to respond to a large-scale problem in our nation's secondary schools. The intended focus of the handbook, on linking instruction and assessment, seemed particularly appropriate, moreover, in view of the fact that most texts for teachers or prospective teachers have focused on either instruction or assessment, rather than on the connections between the two. For example. Gagne' and Briggs'. Principles of Instructional Design (Holt, Rinehart, and Winston. 1979), a highly regarded and widely used text in schools of teacher education, treats issues of learning and teaching extensively, but only 1½ chapters out of 15 in the text deal directly with the assessment of student learning. In a similar manner, Brown's Measuring Classroom Achievement (Holt, Rinehart, and Winston, 1981), a well-respected text, deals thoroughly with topics in measurement, but contains only one chapter out of ten on the use of tests in the instructional process. The anticipated emphasis of the handbook on integrating teaching and testing thus promised to fill a need that rarely was addressed fully in available texts.

The proposed form of the handbook also seemed distinctive. Most texts that deal with classroom instruction or assessment are too long and detailed for practical use by high school teachers. Even those few texts that do give substantial attention to both teaching and testing, for example, Learning and Human Abilities, by Klausmeir (Harper and Row, 1975) and Evaluation to Improve Learning, by Bloom, Madaus, and Hastings (McGraw Hill, 1981), are generally intented to serve as comprehensive texts appropriate for in-depth study. A handbook that would be substantive, but at the same time relatively easy to use in the context of inservice courses and programs for practicing teachers, seemed to be needed.

The development of an inservice program and the preparation of related training manuals to support the Handbook was viewed as essential in light of all that is known about the conditions that foster the adoption of innovations in schools. It is commonly observed that handbooks, texts, articles, and similar materials, no matter how well-conceived or written, rarely in themselves have much impact on teacher practice. Programs must be created that provide teachers with the time, resourses, and assistance needed to translate new ideas and procedures into practice. Research on staff development, as well as our own experience in school improvement efforts, suggest, furthermore, that a variety of roles and responsibilities needs to be carried out in a school district to fully implement an inservice program of the kind pilot tested through this project. It appeared crucial to prepare guides for superintendents, principals and department chairs (or "lead teachers"), as well as participating teachers, that would clarify the tasks each group was to perform.

DEVELOPMENT OF THE HANDBOOK

Overview

From the outset, the Handbook was intended to be both conceptually sophisticated and practical. We had a commitment to draw upon the best available ideas and procedures for connecting instruction and assessment, and to synthesize and present this material in a form that high school teachers would find meaningful.

To help assure that the handbook would be both substantively sound and practical, three different advisory panels were formed in the fall of 1982, soon after the contract was awarded: (1) a technical panel on measurement; (2) a technical panel on instruction; and (3) a panel of practicing high school teachers. Members of these panels are listed in Appendix 1. In addition, Dr. Jason Millman, an eminent authority in testing and the use of test information, was asked to serve as a consultant to the project and, in this capacity, to help integrate recommendations from all three panels about the design and development of the Handbook.

Under the guidance of the advisory panels and Dr. Millman,* the Handbook progressed through five phases of development:

(1) Conceptualization and design (fall 1982 - winter 1983);

(2) Initial draft of selected chapters (spring 1983);

(3) Revision, and preparation of a complete draft (summer-fall 1983);

(4) Pilot testing and subsequent revision (spring - summer 1984); and

(5) Final review and revision (fall 1984 - January 1985)

Issues Faced in Designing and Drafting the Handbook

The most basic and difficult issue faced in designing the handbook concerned its underlying structure. We wanted to find a theme or set of themes around which

---

*Members of the teacher panel met five times between fall of 1982 and fall of 1983. Members of the technical panels did not meet as a group, but rather provided ideas and suggestions by mail and phone. Particularly extensive guidance was furnished by the chairs of the technical panels, Drs. Ronald Hambleton and M. David Merrill, as well as Dr. Millman.

discussions of teaching and testing could be organized. This proved to be a more demanding task than originally anticipated. A number of competing frameworks for organizing the handbook were considered.

One set of organizers discussed was based on the different purposes served by assessment, e.g., to plan instruction in the beginning of a course, to assess students' learning progress, to verify students' mastery over a learning area, etc. If this framework were adopted, the Handbook would focus on designing, selecting, or adapting tests to respond to each of the major types of information needs that teachers have.

Another organizational structure that was discussed centered on instructional models and methods. If this structure were chosen, the handbook would focus on the implications for instructional planning and assessment of different approaches to instruction, either general approaches, like mastery learning, or more specific methods, like simulation, lecture. or discussion. The use of informal methods of assessment, like asking questions in class, or observing students during small group work would need to be described in considerable detail if this framework were adopted.

A third proposal for organizing the Handbook was to focus on the influence that different types of learning goals have on instruction and assessment. What and how well students are expected to learn determines in large part how instruction and assessment are designed. Were this framework adopted, the handbook would include discussions of the kinds and levels of learning outcomes that might be expected from instruction in high school, and guidelines for matching instruction and assessment to these goals.

Each of these candidates for an underlying structure appeared to have merit. No single framework in itself seemed adequate. Many members of both the Teachers' Advisory Panel and of the Technical Panels, for example, made clear that none of these frameworks was likely to permit sufficient discussion of principles for writing technically sound test items and other assessment procedures. Teachers receive such limited

preparation in this area, that many of our advisors held the view that item writing procedures needed to be addressed in some detail.

In the midst of discussions about alternative frameworks, we completed a review of the literature on integrating teaching and testing and a survey of high school teacher's testing practices. The literature review is presented in Appendix 2. The survey instrument is presented and discussed in Appendix 3. These activities provided some assistance in establishing a focus for the handbook. We found through the survey, for example, that teachers expressed much more interest in learning "how to design tests that are matched closely to what I have taught" than other proposed topics concerning the teaching-testing connection, like "how to use test results as a guide for planning remedial or enrichment activities," or "how to score and manage test information with the aid of a micro-computer."

After extensive discussion with the Teachers' Advisory Panel, and the preparation and review by members of both the teacher and experts' panels of several draft chapters of the Handbook, it was decided that the primary conceptual focus for the handbook would be the idea of anchoring teaching, testing, and the use of test information to the learning outcomes expected from instruction, but that the main elements of the other frameworks referred to above also would be introduced. The decision to focus on learning goals as a cornerstone for teaching and testing stemmed largely from the perception that this would permit us to deal more fully with the content of instruction and assessment than would frameworks based on instructional models or the purposes served by tests. This seemed particularly important at the high school level, where so much of what teachers do depends on the kind of content with which they are dealing.

Once the focus of the Handbook was clarified, attention turned to establishing an organizational structure for the Handbook. The most basic organizational decision that was made involved dividing the Handbook into two broad sections. The first, to be called "Foundations," was to introduce the overall idea of integrating teaching and

testing and discuss the various aspects of learning, teaching, and testing that need to be coordinated to achieve this integration. The foundations section would also indicate to the reader that a goal-based approach to linking instruction and assessment would be emphasized in the Handbook.

The second section of the book, it was further decided, would deal with Applications." This would provide guidelines for using the ideas addressed in the foundations section in a deliberate and well-sequenced fashion. Chapters were to deal with formulating learning goals for students, matching instruction and assessment to these goals, assuring that tests meet standards of technical adequacy and are appropriate for one's particular information needs, and so forth. The organizational pattern agreed to is indicated by the table of Contents of the Handbook, which is presented on the following page.

After agreeing on an organizational structure, a couple of months were spent in establishing an approach to classifying different types and levels of learning goals that would serve as an effective guide for instruction, assessment, and the interpretation of assessment results. The most obvious approach to use was Bloom's well-known taxonomy for classifying goals in the cognitive domain. However, after discussion with the Teachers' Panel and in light of our previous experience in applying the taxonomy, we decided that the framework was not satisfactory for our purposes. One reason for this decision was that it was difficult for our teacher advisors, most of whom who had studied the taxonomy before, to differentiate consistently among the six levels of learning described in the taxonomy. Furthermore, the taxonomy was silent about the connections between knowledge (level 1 of the taxonomy) and the general intellectual skills represented in levels 2 through 6. Finally, concern was expressed about the assumption underlying the taxonomy that learning at one level necessarily depended on learning at a lower level. Did "evaluation" invariably involve a higher level of performance than "synthesis"? Bloom's taxonomy, while undoubtedly a major contribution

INTEGRATING TEACHING AND TESTING:

A Handbook for High School Teachers

	Page
Introduction	1
PART I: FOUNDATIONS	3
Chapter 1: What does it mean to integrate teaching and testing?	5
Chapter 2: Dimensions of learning that influence teaching and testing	13
Chapter 3: Dimensions of context that influence teaching and testing	23
Chapter 4: Purposes served by tests	35
Chapter 5: Types of test items and assessment procedures	41
PART II: APPLICATIONS	53
Chapter 6: Formulating expected learning outcomes	55
Chapter 7: Matching teaching to expected learning outcomes	63
Chapter 8: Matching test items and assessment procedures to expected learning outcomes	77
Chapter 9: Assuring quality in tests items and assessment procedures	91
Chapter 10: Preparing, administering and scoring tests	115
Chapter 11: Using test information as feedback to students	135
Chapter 12: Using test information as a guide to instructional decisions	143
Chapter 13: Using test information in grading	159

to educational practice, did not seem entirely appropriate for use in the Handbook.

Various other perspectives on learning goals were studied and discussed, including those developed by Gagne, by Posner and Rudnitsky, and by Merrill. Finally, we decided to adapt Merrill's framework as an anchor for the Handbook, largely because this brought into focus the connections between knowledge and skill, and because it was relatively easy to differentiate among the three levels of accomplishment described in the framework (although we described these levels in different terms than Merrill). More information on our choice of a learning taxonomy is contained in Chapter Two of the Handbook.

Two other content-related issues demanded considerable attention. One had to do with our stance toward informal assessment, for example, the questions teachers ask in class to check students' understanding during a lesson and the unstructured or "spontaneous" observations teachers make of students as they work on individual or group assignments. It was acknowledged that informal assessment is essential to managing instruction and providing feedback to students, but we decided that we would give priority to more formal assessment, that is, assessment based on deliberately structured tasks and supported by well-defined rules for evaluating students' responses. This decision seemed appropriate from the teacher advisors' vantage point in view of the wide experience that teachers have in informal assessment and the fact that numerous inservice programs already have been developed that give considerable attention to this topic.

Another content-related issue that arose concerned the treatment to be given to recent technical advances in domain-specification strategies, standards-setting procedures, item analysis techniques, and related quantitative aspects of testing. From the teacher advisors' point of view, none of these technical developments had any direct relevance for classroom assessment, largely because they required a level of training and time-commitment beyond that of virtually any practitioner. Several members of

the technical panels urged inclusion of some of this technical material in a simplified form. We decided not to discuss complex technical advances in any depth, but to include a list of resources at the end of each chapter that the reader could consult for extended treatments of these topics. In retrospect, we may have erred on the side of excluding information on basic quantitative aspects of testing, such as item analysis procedures, but, given the intended audience for the Handbook and the context in which it is to be used, the decision to keep discussions of technical measurement topics to a minimum seemed warranted.

Finally, many more specific issues were encountered in the process of writing individual chapters. For example, when writing Chapter 6, on formulating learning goals for students, members of the Teacher's Advisory Panel indicated strong dissatisfaction with traditional prescriptions for formulating objectives, which require teachers to specify not only the knowledge and skills students are to acquire or apply, but the conditions under which student learning will be assessed and the criteria against which student performance on assessment tasks is to be evaluated. Drs. Millman and Hambleton also suggested that including performance criteria in goal statements was impractical. and not necessarily desirable. Several drafts of Chapter 6 were required in order to establish guidelines for writing goals that appeared both sound and feasible.

The issue of how best to write learning goals was merely one of many that were dealt with on a chapter-by-chapter basis, and that often led us to revise substantially what we had initially written. In fact, each of the chapters in the Handbook was revised at least once, and in many cases, two or three times, to assure that its contents were treated appropriately. Critiques provided by Drs. Millman and Hambleton were particularly helpful in this process.

Distinctive Characteristics of the Final Product

The distinctive characteristics of the Handbook have been suggested in the previous section. These characteristics are summarized and made explicit below.

(1) The Handbook sets forth in reasonable detail a goal-based approach to integrating teaching and testing. A conceptual perspective on the nature of learning goals is introduced in Chapter 2 and used as a basis for organizing Chapter 6, on formulating learning goals, Chapter 7, on matching instruction to goals, Chapter 8, on matching assessment to goals, and on major sections of Chapters 10, 11, 12, and 13. We believe the idea of using learning goals as a basis for linking instruction, assessment, and the use of assessment information is both theoretically sound and responsive to the general trend in education to require both teachers and school districts to make public the learning outcomes students are expected to achieve in courses and programs and to report evidence on outcome attainment to students, parents, and members of the community.

(2) Having been developed under the guidance of 12 high school teachers, the Handbook reflects sensitivity to contextual factors that influence teacher practice. Chapter 3, for example, includes a description of the influence that student characterisitics have on teaching and testing, and the implications that different instructional models have for these processes. More generally, the guidelines and numerous illustrations contained in the document have been carefully developed to reflect high, but attainable standards for teachers' professional performance.

(3) The Handbook has been pilot tested in the context of a school-based inservice program, and carefully reviewed by nationally-known technical consultants. Both its content and form have been judged to be of good quality and appropriate for the intended audience.

DEVELOPMENT AND PILOT TESTING OF A STAFF DEVELOPMENT PROGRAM BASED ON THE HANDBOOK

During the fall of 1983, a small-scale, school-based inservice program was developed in collaboration with four members of the Teachers' Advisory Panel. The Panel as a whole had completed its work in guiding the design of the Handbook, but these four members expressed a special interest in the project and volunteered to assist in developing and pilot testing an inservice program around the Handbook. We refer to these teachers as "lead teachers."

The inservice program was intended to provide an opportunity for lead teachers to work with colleagues for two full days in their home districts on the integration of teaching and testing. More specifically, the program was designed to permit participants to review and discuss the ideas and methods contained in the Handbook and to use this content as an aid in developing an instructional unit, appropriate for their setting and needs, that made clear connections between instruction and assessment. More detailed information about the inservice program is offered below.

Purposes to be Served by the Pilot Test

The pilot test of the Handbook and the related inservice program was designed to serve the following purposes:

- To assess high school teachers' reactions to the Handbook as a resource for integrating teaching and testing;
- To assess high school teachers' reactions to the inservice program developed around the Handbook.
- To assess the reactions of lead teachers to the inservice program and the role they carried out in it; and,
- To serve as a basis for the design of instruments for observing teaching-testing practices, and for assessing their effects on student attitudes.

Sample

Schools. Four comprehensive public high schools participated in the study. One school was in southwestern Washington. The other three schools were in western Oregon. The schools ranged in size from 460 students to 1,800 students. Three schools served students at grades 10 to 12; one spanned grades 9-12.

Lead Teachers. A lead teacher from each school served as facilitator of the inservice program. The lead teachers met for two days during the fall and winter of 1983 to prepare for their roles as facilitators. This preparation involved shaping agendas for the work sessions, reviewing relevant content, and drafting illustrative teaching and testing material for a particular unit of instruction. In addition, one of these teachers met individually with project staff in early January to discuss content and procedures about which he had questions and to refine the illustrative teaching-testing materials he had developed. The lead teachers represented the subject areas of science, social studies, and foreign languages.

Regular Teachers. Two teachers from each of the four schools participated with the lead teacher in the inservice program. In each school, these teachers were selected by the lead teacher and the principal because of their interest in the proposed content of the program. Participation was voluntary. These teachers represented the following subject areas: science, art, social studies, French, mathematics, and English. Unfortunately, due to illness one teacher from each of two schools was unable to attend the second day of the two-day program (the program is described later in this report). The total number of regular teachers who participated fully in the program was thus 6.

Project Staff. A member of the project staff participated in each of the two all-day inservice sessions in the four schools. His role was to provide a context for the Handbook and the inservice program; to clarify issues of content that arose from teachers' reading of the Handbook and their effort to translate ideas from the Handbook into teaching-testing material; and to support the lead teacher in any way that seemed

appropriate as the lead teacher guided his colleagues through the designated activities.

Principals. In three of the four schools, project staff met personally with the principal in the late fall or early winter to discuss the inservice program. All principals had been informed about the program in the spring of 1983, but a face-to-face meeting was considered desirable to firm plans and to assure the principals' support. In one school, the principal gave clear verbal support by phone, and did not feel a need for a personal meeting. All principals received the draft Handbook and a synopsis of the inservice program.

Principals had no role to play in the field test other than to communicate their support of the program to teachers who were participating in it. One principal, however, took the initiative to attend about 15 minutes of the first day of the inservice session. He suggested that he took this step to increase his understanding of the program and to actively demonstrate his interest in the work being done.

Expected Outcomes

The original design for the inservice program set forth a wide range of expectations. Teachers were expected to read and discuss a large portion of the Handbook, and to develop an instructional unit that included not only statements of expected learning outcomes, tests, and descriptions of anticipated uses of test information, but information on teaching methods, and informal, day-to-day assessment tactics. During the first work session at the first school to field test the program, it became apparent that the focus of the work sessions needed to be narrowed somewhat. Both of the regular teachers needed more time than was allocated to formulate expected learning outcomes according to the framework proposed in the Handbook. (An English teacher in the session had particular difficulty with this task because, by her own account, she was unaccustomed to specifying student learning goals or objectives.) The teachers in this first session also indicated that they were most interested in thinking about ways of strengthening their approach to testing outcome attainment. The original

plan called for a description of instructional strategies and informal assessment procedures, but the teachers did not assign these tasks as high a priority as test development. They noted that they had received training in instructional strategies, but had little formal background in testing. Therefore, given the time constraints, they suggested that they focus on instructional strategies only to the extent that they had concerns about them. This seemed reasonable to the lead teacher and the project staff member, each of whom pointed out that issues of instruction could be addressed later in the context of a discussion of the instructional implications of test results.

Based on the first two or three work sessions, in which various modifications in scheduling and content coverage were tried and evaluated, a more narrowly defined set of expectations was established that seemed more feasible to achieve in the time available. The revised set of tasks that participants in the program were expected to complete is presented in Display 1.

Training Procedures

The inservice program was intended to provide an opportunity for teachers to work both together and individually on the integration of teaching and testing. More specifically, the program was designed to permit participants to review and discuss the ideas and methods contained in the Handbook and to use this content as an aid in carrying out the tasks identified in Display 1.

Two full-day work sessions were allocated for the program in each building (a total of eight sessions were thus held). These took place between mid-January and early March, 1984. These sessions began at 8:30 and adjourned at 3:30. For each of the major tasks to be accomplished (Display 1), time was given to review relevant chapters of the Handbook and to discuss issues raised through the reading. Each participant received the Handbook at least 3 weeks prior to the sessions, so some familiarity with its content and organization was assumed. Following a discussion of relevant content, participants had an opportunity to work on a designated task.

DISPLAY I

PILOT TEST OF THE HANDBOOK AND RELATED STAFF DEVELOPMENT PROGRAM

January, March, 1984

Tasks to be Completed Through the Two All-Day Work Sessions, Supported by About Two Hours of Time Outside of the Work Sessions

1. Write learning goals and objectives for an instructional unit that you plan on teaching between March 1 and May 1 of this year. Classify these goals and objectives according to the framework provided, and indicate the relative emphasis each is to receive for purposes of instruction and assessment.

2. Develop a unit test or comparable assessment procedure that matches unit goals and objectives and the emphasis that is to be placed on each. Use the guidelines provided in chapters 9 and 10 of the Handbook to help asure that your test meets standards of technical adequacy.

3. Set a standard of acceptable test performance for the class as a whole. (You may also set separate standards for individual students. If so. provide a rationale for this.) Indicate the type of options you will pursue if these standards are not met. Specify how many class days, if any, you are willing to devote to helping students correct learning deficiencies revealed through the unit test. Also specify any provisions that will be made for retesting students after corrective instruction has been completed. Finally, describe what type of enriching instruction, if any, will be given to high-achieving students, while others are receiving corrective instruction.

4. Describe procedures for reporting and reviewing test results with the class as a whole, and for giving feedback to individual students on their test performance.

5. Specify how test results will be used in calculating a grade for the marking period, quarter, or semester.

Participants worked on the tasks individually, but then shared work products with their colleagues. After a review and discussion of the products, some time was allotted for individuals to refine their work, although participants were expected to fine-tune material outside the context of the work sessions.

The lead teacher in each building and the project staff member (the same staff person attended all work sessions) guided the sessions. The kind of partnership that emerged between lead teacher and project staff varied across sites. In two of the sites, the lead teachers were accomplished leaders of inservice programs in their district and had also invested considerable time in preparing illustrative units intended to highlight the link between instruction and assessment. In these two schools, both the project staff member and the lead teacher worked together to explain expectations, clarify content, and respond to concerns about the applicability of ideas and procedures to particular topics and classes.

Two of the lead teachers assumed a less active posture in the work sessions. They deferred more to the project staff member both with respect to explaining content and commenting on teachers' work.

Reactions of Participants to the Handbook and to the Inservice Program

The main tool used to assess participants' reactions to the Handbook and the inservice program was an 11-item questionnaire. This questionnaire was given to the regular teachers at the end of the second day of the work sessions. Two of the eight regular teachers were absent for the second day and did not receive the questionnaire. Informal phone conversations revealed that these teachers had positive attitudes toward both the Handbook and the first day of the program, but no formal assessment of their reactions was conducted.

A report of responses to each of the eleven items on the questionnaire on the part of the six teachers who completed it is presented below, followed, when appropriate, by a brief commentary. As will be seen, the general response to both the Handbook and

the inservice program was quite positive.

Question 1: About how much time, if any, did you spend reading the Handbook **before** the work sessions?

Mean response: 1 hour, 18 minutes
Range: 30 minutes to 2 hours, 45 minutes

Teachers had been asked to read the first several chapters in the Handbook and chapter 6, on formulating learning goals. The reading was intended to convey the overall approach to integrating teaching and testing with which the participants would be dealing, and to prepare them specifically to write learning goals for "trial" instructional units. The personal impression of the staff member attending the sessions was that these self-reports overstated the actual time invested. When asked to comment on the framework for thinking about learning goals described in the Handbook, for example, few of the participants could clearly recall its basic features. Perhaps this was because insufficient structure was provided to guide the teachers' reading.

Question 2: About how much time, if any, did you spend reading the Handbook **between** the first and second day of the work sessions?

Mean response: 1 hour, 35 minutes
Range: 45 minutes to 4.5 hours

Teachers were expected to read chapters on the use of assessment information (Chapters 11-13) between the first and second sessions. Most seemed to have done this and were able to point out at least one or two ideas they found useful, or at least noteworthy.

Question 3: Please describe in a brief paragraph how you used the Handbook **during** the work sessions.

Representative teacher responses to question 3 are presented below. (They have been paraphrased slightly for ease of reading).

Teacher 1: I used it primarily to give examples of things that I might try.

Teacher 5: During the first day, the handbook clarified vocabulary and content questions. During the second day it provided actual help for test writing and evaluation.

Teacher 6: For clarification - so I could be sure I fulfilled my responsibilities - also for interest - I was intrigued by the concepts.

Question 4: Which part or parts of the Handbook did you find most helpful?

Teacher 1: The parts of the handbook involving designing tests - particularly the "memory" and "use" type of questions. The methods for evaluating by observation were very useful.

Teacher 2: The areas on writing objectives and writing test questions.

Teacher 3: Chapter 7 was a good summary of techniques used to teach. Chapter 9 - the most useful - specific examples of types of test items. For me, all the theorizing in the book could have been reduced and extra emphasis placed here. Chapter 10 also contained good information on testing.

Teacher 4: Chapter 9 was probably the most helpful. I also looked closely at chapter 12, about what to do with students who fail exams.

Teacher 5: Parts dealing with test item writing, reviewing tests, and those students who don't grasp materials. I found the examples in each section to be extremely helpful!!

Teacher 6: I found several parts helpful - level of performance and its tie with Bloom's taxonomy, securing motivation in a low interest student, assessing why a student is failing are just three examples. They were helpful in that they offered an intelligent, practical approach to the problem. It is very difficult to answer this question as I found so much that was helpful.

Question 5: Which part or parts of the Handbook did you find least helpful? How do you think these parts could be improved?

Interestingly, three of the six teachers left this blank. One who did respond said at first she thought there would be too much to read, but then found the assignments acceptable. She concluded, "I have no suggestions to make."

Another respondent said the comparison between the framework for viewing learning goals presented in the Handbook and Bloom's taxonomy was least helpful because he "...didn't grasp as much of this as I'd like." Finally, an art teacher stated that the section on writing objective test items was not very relevant to his class, though he indicated that the material on product-related assessment was pertinent.

Question 6: In your view, how much value has the Handbook been as an overall resource for planning instruction, developing interests, and using test information?

Mean response: 3.1 on a 4 point scale in which 1 = no value and 4 = a great deal of value

Range: 2 to 4

Question 7: In what ways, if any, do you feel you benefited from the work sessions on integrating teaching and testing?

Teacher 1: The assistance of the instructor(s) was very beneficial in helping me design my tests toward what I teach, and my teaching more directly tied to what I want students to learn.

Teacher 2: The time to write goals and tests is very limited. This time was valuable for that. I really felt that having a resource person here was very helpful.

Teacher 3: I feel that I benefited from some of the ideas of ways to test for different learnings - giving different emphasis to different styles. Since I use a text-related standard test, I think I am better able to evaluate good and bad items in the tests.

Teacher 4: I am more aware of the expectations that I have of students - also the limitations.

Teacher 5: It gave me the opportunity for 2 full days of work. Glen's guidance and the time were the most valuable.

Teacher 6: I gained a new perspective which enabled me to improve my testing. With an adequate background in the subject area. and a few years teaching experience, it is easy to get in a "rut" that repeats weaknesses. This session influences that type of weakness.

Question 8: Was there anything about the work sessions that was disappointing or negative? How, if at all, could these aspects of the work sessions be improved?

Teacher 1: No

Teacher 2: I wish we could have been away from the school area. Too many distractions.

Teacher 3: I would have like a more structured approach. I think there was too much theorizing. I also find it very difficult to try to formulate ideas in a setting where I almost felt like I was being "watched." I prefer to work alone - I would have preferred to concentrate on my own goals - alone.

Teacher 4: It was a one-on-one situation, so the immediate response was great. However, time was of the essence - needed more time - another day.

Teacher 5: The first day's setting was not a working setting for me. Too noisy, and not enough time to work alone.

Teacher 6: I can offer no suggestions.

The issues of setting and time were apparent to the project staff member who attended the sessions. Four individuals working together in a relatively small room for two six-hour days represented an intense work environment. And, even though the

number of tasks had been reduced after the first inservice session, it was difficult to achieve all the outcomes desired in the time allotted.

Question 9: By the end of the second day of the work sessions, how far along were you toward completing the instructional unit call for in the project?

Mean response: 3.2 on a four-point scale in which 1 = "I didn't start the unit" and 4 = "I completed nearly all of the unit

Range: 2 to 4

Question 10: How, if at all, was the approach to unit development introduced in the work sessions different from the approach you normally follow? Was there anything about this approach that you especially liked or disliked?

Teacher 1: Planning the unit around specific outcomes and planning the test based more closely on what students were supposed to learn gave a truer picture of what actually was and was not learned.

Teacher 2: The approach was low key, and I wasn't sure about the direction we were going at first, but as we worked along I felt more organized.

Teacher 3: I don't want to imply that I don't organize my thoughts before teaching, but teaching elementary levels of French involves so many minute components that the idea of having to write goals and objectives for each one is overwhelming. Therefore, some of the writing of goals is a waste of time to me.

Teacher 4: I normally teach in units. It was ok.

Teacher 5: The approach was more detailed, with a definite beginning and ending. Things are usually done day to day for me. I enjoy this approach and will attempt to use it regularly.

Teacher 6: It was a more "planned" approach. I liked feeling that I had produced a good test - it gave me more enthusiasm and energy as I taught the unit..

Question 11: How much help were the work sessions to you in strengthening the links between your teaching and testing practices?

Mean response: 3.1 on a 4-point scale, in which 1 = no help and 4 = a great deal of help.

Range: 2 - 4

On the whole, then, the reactions to both the Handbook and the inservice program certainly were encouraging. Judging from responses to the questionnaire it appeared that the work sessions assisted teachers in thinking about instruction and assessment more systematically. However, no objective evidence on the effects of the program

on teacher practice or on students was gathered during the pilot test. As is indicated in the last part of the report, however, evidence of this kind on the effects of the inservice program is currently being collected as part of a related project supported by the NIE.

Reactions of Lead Teachers

Reactions of lead techers to the inservice project were assessed informally after each of the school based sessions, and then again in June at a luncheon meeting.

With respect to the content of the program, these teachers indicated that they refined a number of their ideas and practices regarding instruction and assessment based on their involvement in the project. Here are some representative comments on this topic. (They are paraphrased slightly for ease of reading).

> **Teacher A:** The project has made me think a lot more about what I want to focus on and then test. I never thought much about giving different weights to test items, for example, based on the emphasis I've given particular objectives in class. Also, the idea of paying more attention to what test results imply about next steps to take in instruction stays with me...I've been trying some new ways of helping kids to figure out why they got some questions wrong on a test.
>
> **Teacher B:** Letting students in on what the learning goals are seems important. I do this in a clearer, more direct way now. I also look at the goals when making tests, whereas before I had them more in the back of my mind... I think thinking about goals more tends to push you into higher levels of learning... I also am giving more quizzes and doing more checks during units on how students are grasping the material.

With respect to the processes used in the program, the lead teachers generally felt satisfied with the work sessions and their role in them, but expressed these suggestions: (These, too, are paraphrased slightly).

> **Teacher A:** "Have the inservice earlier in the year, maybe October or early November. This will help people think more about the course as a whole and also let them try out new ideas over a longer time period.
>
> **Teacher B:** If possible add a day or two to the inservice so that topics could be dealt with more fully.
>
> **Teacher C:** Consider having teachers from the same department meet together. While there is some value in having people from various fields get together, teachers from the same department would be able to deal more effectively with substantive issues.

Teacher D: Build in more "accountability" between lead teachers and their colleagues. Lead teachers need to follow-up on what their colleagues do after the work sessions and give support and assistance as needed. This idea was discussed in the project, but a system wasn't set up to help it get done.

This last point seemed particularly important, though no easy resolution of the issue seemed possible without greater involvement of school administrators and perhaps more released time, or more flexible teaching schedules, for teachers participating in the project. As with many inservice efforts aimed primarily at the classroom teacher. the implications of the program for school policy and management issues were large. Many of these issues are addressed in the implementation manuals, which are described in the next part of the report.

Use of the Pilot Test as a Basis for Designing Research-Related Measures

During the pilot test a member of the staff of the project began work on a related project sponsored by the Center for Educational Policy and Management (CEPM) at the University of Oregon, with funds provided by the NIE. This project was intended to serve as a sequel to the handbook-development project. Its aim was to refine the staff development program being pilot tested and to conduct research on the effects of the program on teacher practice and student attitudes. The CEPM project also was intended to yield information on administrative policies and procedures that affect classroom practices. The CEPM project is described more fully in the last section of this report.

The pilot test of the staff development program described in the preceeding pages provided a context within which to devlop the instruments called for in the CEPM project. These instruments included (a) procedures for observing a teacher's approach to giving feedback to the class on its test performance; (b) a guide for interviewing teachers before and after the observation to obtain supplementary information on test-related practices; and (c) a questionnaire for students on their attitudes toward coursework, testing, and grading, and toward themselves as learners.

Each of the four lead teachers involved in the pilot test kindly agreed to allow project staff to observe their classes on days when tests were being returned and to videotape some of these lessons. This helped immeasurably in the development of the classroom observation instrument referred to above. Two of the lead teachers, along with other participants in the inservice program and several teachers from schools not previously involved in the project, agreed to pilot test the teacher and student questionnaires prepared for the CEPM project. The pilot test of the inservice program thus provided a basis for developing the measures needed to carry out research on the effects of this kind of program on teachers and students.

PREPARATION OF STAFF DEVELOPMENT MANUALS FOR TEACHERS, LEAD TEACHERS AND PRINCIPALS, AND SUPERINTENDENTS

Experience gained through the pilot test of the Handbook and the 2-day inservice program on its use provided the basis for the preparation of these manuals. They are responsive to the expressed needs of teachers who engaged in the inservice program; they are responsive to the experience of the lead teachers who conducted the program, and the suggestion of principals who had to approve and support the program in their buildings; and they anticipate the essential role that Superintendents play in the decision of a school or district to initiate a staff development program for its teachers that is intended to change the nature of school practice and requires the expenditure of considerable time and energy to do so.

In combination these manuals are intended to provide a "stand alone" training system that districts will be able to purchase at low cost and implement with their own personnel. The training system rests on the central assumption that with adequate resources a principal, and a staff member selected by the principal to serve as a "lead teacher" for a department, can function effectively as a training team for faculty within the teacher's department. The instructional resources needed to conduct the training program are carried in both the **Resource Guide for Teachers** and the **Planning Guide for Lead Teachers and Principals.** The information needed to explain the program to a Superintendent, and to convince him or her of the value to be gained through it in relation to the costs involved, is carried in the Program Guide For Superintendents.

This section of the final report contains brief descriptions of the content and organization of those manuals, and how they have been developed. It also contains a description of the field trials and refinement process that is planned for the manuals. As the manuals presently stand, only parts of the first half of the **Resource Guide for Teachers** (the part of the Guide that contains a two-day BASIC training program) has been used and reviewed by teachers.

The Resource Guide for Teachers

The staff development program is divided into a two-day BASIC training program and a three-day ADVANCED training program. The BASIC program focuses on the integration of teaching and testing practices in the context of **instructional units**, and is linked tightly to the content of the Handbook developed through the project. It also parallels closely the two-day inservice program that is described in the previous section of this report, and the two-day training program that constitutes the treatment condition in an ongoing research study funded through the Center for Educational Policy and Management at the University of Oregon (see the last section of this report).

The ADVANCED program focuses on the integration of teaching and testing practices in the context of a **course of study**, and addresses the issue of how faculty within a department articulate learning goals across courses within a **program of study**. The advanced training program also ties closely to the content of the Handbook, but it is not limited to this content. Appendices included in the Resource Guide, and information brought to the program by the principal, supplement the Handbook in topic areas that go beyond it.

In combination the Handbook, the Resource Guide and the information that lead teachers and principals are able to bring to the training sessions are intended to provide all of the instructional resources needed to prepare teachers to effectively integrate teaching and testing practices in the courses they teach. An outline of the content of the Resource Guide is provided in DISPLAY 2. For each day outlined in this display, the Guide contains a suggested work plan for the day and an instructional unit for each of the topics listed.

The Planning Guide for Lead Teachers and Principals

The Planning Guide parallels the Resource Guide in its organization and content. There is an introductory section that outlines the roles and responsibilities of lead teachers and principals in the training program, and the philosophy underlying the

DISPLAY 2

TABLE OF CONTENTS FROM THE RESOURCE GUIDE FOR TEACHERS

Page

A Note to Teachers .....

INTRODUCTION .....

PART I. THE BASIC TRAINING PROGRAM .....

Overview .....

Outcomes to be achieved .....

Day 1: Establishing learning goals for a UNIT of instruction, designing an assessment plan, and developing measures to assess student mastery of the learning outcomes desired .....

Day 2: Setting standards for goal mastery, deciding how performance standards are to be applied to slow and fast learning students, and giving feedback to students on test performance that highlights what has and has not been mastered .....

PART II. THE ADVANCED TRAINING PROGRAM .....

Overview .....

Outcomes to be achieved .....

Day 3: Reviewing the effects on student learning of work accomplished in the BASIC training program, refining the framework used in the basic program for thinking about learning goals, and articulating UNIT goals with COURSE and PROGRAM goals .....

Day 4: Designing an assessment plan for a course of instruction, linking grading standards for a course with performance standards for goal attainment, and reaching departmental and school-wide agreement on grading standards, including standards for slow and fast learning students .....

Day 5: Using information about goal attainment to plan a COURSE of instruction, guidelines for managing goal-based instruction with students who are working at expected levels of accomplishment, and guidelines for students who are working toward learning goals that deviate appreciably from expected levels of accomplishment .....

APPENDICES

A. An Introduction to Mastery Learning
B. Notes on the Articulation of PROGRAM, COURSE and UNIT goals
C. Illustrative Learning Goals for a Course of Instruction

training program generally. There also is a discussion of the key role that a superintendent needs to play in reviewing and approving the program, and in either lending support or giving leadership to its implementation.

The core of the Planning Guide, however, is the NOTES AND RESOURCES for lead teachers and principals that accompany each day of the training program. These parallel and build on the instructional resources presented in the teacher's Resource Guide, and are intended to provide the wherewithall the lead teachers and principals need to function effectively as facilitators of the training program. In this regard, the recommendation is made in the Planning Guide that the lead teacher/principal teams who are to facilitate the training program within a district engage in at least a week of intensive preparation/study before the training program is undertaken. It is recommended that this be done under the tutelage of someone in the district or a nearby college who is knowledgeable about tests and current testing in the schools.

The Program Guide for Superintendents

The Program Guide serves a very different function than the two manuals that have been described previously, though it is viewed as being complementary to the other manuals and essential to the program as a whole. The purpose of the Program Guide is to inform the superintendent, other administrators and members of the local Board of Education about the benefits to be gained by integrating teaching and testing practices within a district, and to orient them to the training program as a means of doing so. An overview also is provided as to what the integration of teaching and testing practices means operationally, the implications of such practices for staff development and the cost of implementing the training program. The preparation of the Program Guide is founded on the assumption that a change in the nature of schooling of the kind that comes with the effective integration of teaching and testing is one that must be considered carefully by everyone concerned before a decision to make such a change is made. It also rests on the fact that in smaller districts resource expenditures

of the kind required to implement the training program are always subject to the review and approval of the superintendent.

Plans for Field Testing and Refining the Staff Development Manuals

As indicated previously none of the manuals have been field tested in their present form. Four of the six instructional units contained in the 2-day BASIC training program were used in the pilot test of the Handbook, and currently are being used in the CEPM sponsored experiment described in the next section of this report, but that is the extent to which they have been subjected to the realities of field use. An extensive program of field review and testing is planned for the months ahead, however, and revisions/refinements will be made in the manuals on the basis of them. The refined versions of the manuals will be submitted to the NIE in conjunction with the final report on the CEPM sponsored experiment in this area that currently is underway. The field review and testing program that is planned is outlined in DISPLAY 3.

DISPLAY 3

PLANS FOR FIELD TESTING AND REFINING THE STAFF DEVELOPMENT MANUALS

The Resource Guide for Teachers

1. A critique by teachers and lead teachers participating in the CEPM experiment (Winter and Spring, 1985).

2. A critique by members of the Valley Education Consortium Planning Council, which includes five secondary principals, four elementary principals, and four district curriculum coordinators (Spring, 1985).

3. A critique by members of the Valley Education Consortium work group on Program Evaluation, Troubleshooting and Improvement, which includes four superintendents, four curriculum coordinators and for purposes of the review four teachers who have been nominated to serve as lead teachers in the superintendents' districts (Spring, 1985).

4. A critique by lead teacher and principal teams within VEC districts who have completed a week of training with the Resource Guide and Planning Guide in preparation for implementing the training program in their schools during the 1985-86 school year (late Summer, 1985).

5. A critique by teachers, lead teachers and principals who participated in the training program in VEC districts (late Fall, 1985).

The Planning Guide For Lead Teachers And Principals

1. A critique by lead teachers and principals participating in the CEPM experiment (Winter and Spring, 1985).

2. A critique by members of the Valley Education Consortium Planning Council (Spring, 1985).

3. A critique by members of the Valley Education Consortium work group on Program Evaluation, Troubleshooting and Improvement (Spring, 1985).

4. A critique by lead teacher and principal teams within VEC districts who have completed a week of training with the Resource Guide and Planning Guide in preparation for implementing the training program in their schools during the 1985-86 school year (late Summer, 1985).

5. A critique by teachers, lead teachers and principals who participated in the training program in VEC districts (late Fall, 1985).

The Program Guide For Superintendents

1. A critique by members of the Valley Education Consortium work group on Program Evaluation, Troubleshooting and Improvement (Spring, 1985).

2. A critique by lead teacher and principal teams within VEC districts who have completed a week of training with the Resource Guide and Planning Guide in preparation for implementing the training program in their schools during the 1985-86 school year (late Summer, 1985).

3. A critique by superintendents of the VEC districts in which the training program has been implemented (late Fall, 1985).

RELATION TO CURRENT WORK SUPPORTED BY THE NIE

As indicated earlier in the report, the NIE has provided funding for a research project, sponsored by the Center for Educational Policy and Management (CEPM) at the University of Oregon, that builds upon the handbook development project. The CEPM Project extends the original work in two main prospects. First, it is intended to produce more information about school and department policies that influence how teachers carry out instruction and assessment in their classrooms. We knew from our experience in the Handbook-development project and our work with schools generally that the curriculum adopted by a school, the textbooks in use, the approach to teacher supervision that is taken, the availability of test-item banks and test-scoring machines, and other factors that lie outside the control of the individual teachers influence what teachers do in the classroom. The CEPM Project has provided an opportunity to explore connections between aspects of the school context and classroom practices.

The CEPM Project also has provided an opportunity to refine the inservice program that was pilot tested in early 1984 and to conduct research on the effects of the refined program on teacher practices and student attitudes. The research follows an experimental design in which science teachers from five high schools participate in an inservice program, and science teachers from five other high schools serve as a control group. Data on teacher practices and student attitudes are collected on a pre and post-treatment basis. The progress that is being made on the CEPM study is reported periodically to the NIE.

REFERENCES

Burry, J., Catterall, J., Choppin, B., & Dorr-Bremme, D. Testing in the nation's schools and districts: How much? What kinds? To what ends? At what costs? Report No. 194. Los Angeles, CA: Center for the Study of Evaluation, University of California, 1982.

Carter, K. Do teachers understand principles for writing tests? Journal of Teacher Education, 1984, November-December, 57-60.

Rudman, H., Kelly, T., Wanous, D., Mehrens, W., Clark, C., & Porter, A. Integrating assessment with instruction: A review (1922-1980). East Lansing, MI: Institute for Research on Teaching, Michigan State University, 1980.

Sirotnik, K. A. What you see is what you get - consistency, persistency, and mediocrity in classrooms. Harvard Educational Review, 1983, 53, 16-31.

Stiggins, R., & Bridgeford, N. Final research report on the role, nature, and quality of classroom performance assessment. Portland, OR: Northwest Regional Educational Laboratory, November, 1982.

APPENDIX 1

Members of the Advisory Panels the Guided the Development of the Handbook

Members of the Advisory Panels that Guided the Development of the Handbook

Technical Panel on Teaching

David Merrill (Chair)
University of Southern California

Jeri Benson
University of Maryland

Harvey Black
Brigham Young University

Bennie Lowery
San Diego School District

Charles Reigeluth
Syracuse University

Teacher Advisory Panel

Alice Befus
Willamina HS, OR

Jim Ferguson
Cascade HS, OR

Dave Hagfeldt
South Albany HS, OR

Jan Jobe
Central HS, OR

Diane John
Parkrose HS, OR

Darrell Johnston
Dayton HS, OR

Larry Johnston
Woodburn HS, OR

Bob Kenyon
Dallas HS, OR

Al Ko
Ft. Vancouver HS, WA

Jon Lewis
McMinnville HS, OR

Mike McVay
Glencoe HS, OR

Lou Mueller
Sunset HS, OR

Technical Panel on Testing

Ronald Hambleton (Chair)
University of Massachusettes

Eva Baker
University of California
Los Angeles

Ronald Berk
Johns Hopkins University

Tom Haladyna
Americal College
Testing Program

Joan Herman
University of California
Los Angeles

Paul Williams
Maryland Department
of Education

APPENDIX 2

A Review of the Literature on Integrating Testing and Teaching

NIE Project #400-82-0013
Developing a Handbook for Integrating
Testing and Teaching

INTEGRATING TESTING AND TEACHING:

AN UPDATED REVIEW OF THE LITERATURE

Glen Fielding

Teaching Research Division
Oregon State System of Higher Education
Monmouth, Oregon

January, 1983

INTEGRATING TESTING AND TEACHING: AN UPDATED REVIEW OF THE LITERATURE

In recent years, educational researchers and practitioners have paid increasing attention to the relationship between testing and teaching. The growing interest in this topic is signaled by a number of conferences and publications, most notably the review of the literature on integrating assessment with instruction prepared by Herbert Rudman and associates (1980) at the Institute for Research on Teaching; the conferences sponsored by the National Institute of Education in 1978 on research on testing, teaching and learning, and on achievement testing and basic skills; the monograph series, New Directions for Testing and Measurement, published by Jossey Bass; the forthcoming spring issue of the Journal of Educational Measurement, which is devoted to the theme of linking testing and instruction; a series of reports on the use of standardized tests in American schools, produced by a team of researchers at Carnegie Mellon University and the University of Pittsburgh (Resnick, 1981); and a comprehensive survey of test usage in schools and districts completed by investigators at the Center for the Study of Evaluation at the University of California at Los Angeles (Burry et al., 1982).

The present review has been prepared as backdrop to the development of a handbook intended to help teachers better integrate assessment and instruction. Special attention is paid in the review to information on the testing-teaching linkage at the high school level, since the handbook is being designed for use by teachers at grades 9-12.

The review is based on findings and conclusions reported in the conference proceedings and publications referred to earlier, and on recent studies not cited in these documents, in relation to the following questions:

(1) What types of tests and assessment procedures do teachers typically use, and for what purposes?

(2) What are teachers' attitudes toward and concerns about testing?

(3) What do teachers generally know about testing, and what aspects of testing, or test usage would they like to know more about?

(4) What advances in the technology of testing, if any, are likely to strengthen the link between testing and teaching?

(5) What specific approaches to integrating testing and teaching have been developed, and how applicable are they to regular high school classrooms?

These questions are addressed in the review in a somewhat distinctive fashion. A "table" that summarizes research and development work pertaining to each question is provided. These tables represent the substantive core of the review, and are referred to throughout the discussion.

1. <u>What types of tests and assessment procedures do teachers typically use, and for what purposes?</u>

At both the elementary and secondary school levels, teachers report using a wide range of tests and assessment procedures, from standardized tests, to district developed tests, to tests included in curriculum materials, to self-constructed tests, quizzes and assignments, to classroom observations (Burry et al., 1982). Teachers use tests administered by a school district most frequently for planning and grouping decisions at the beginning of the year, though results from these tests also appear to play some part in grouping decisions made during the year and decisions about students' report card grades (Burry et al., 1982). Classroom observation, and in some cases teacher-made tests, are used at the beginning of the year for grouping decisions, for verifying test information, and for developing an initial awareness of students' learning characteristics and attitudes. Observations and teacher-made tests clearly play a much more central role in teachers' daily and weekly planning, and in grading, than do standardized or district-developed tests. Classroom observations also help guide the moment-to-moment, or "in flight," decisions that teachers make while interacting with students.

Patterns of test usage are quite similar at the elementary and secondary school

levels, with two notable exceptions. Secondary teachers use tests included in curriculum materials less frequently than do elementary teachers. For example, 68 per cent of the high school teachers surveyed in research done by Burry and associates (1982) indicated that results from tests accompanying curriculum materials were important in determining students' report card grades, whereas 93 per cent of the elementary school teachers surveyed found these tests important in grading. Secondary teachers also use standardized tests less frequently than do elementary teachers, though the extent of standardized test use in secondary schools varies considerably by teacher subject matter specialization (Burry et al., 1982).

More detailed summaries of teachers' use of standardized tests and classroom observations are provided in Tables 1 and 2, respectively. Detailed summaries have been prepared in reference to these sources of student information because they have been studied much more extensively than other aspects of classroom testing, such as the use of teacher-made tests.

2. What are teachers' attitudes toward and concerns about testing?

Most research on teachers' attitudes and concerns about testing has focused on standardized and state or district-developed tests. By and large teachers appear to see some value in these tests, so long as they are used in conjunction with other sources of information on students.

More specific research findings on teacher's attitudes toward and concerns about testing are presented in Table 3. It is noteworthy that teachers' attitude vary according to the level of knowledge they have about tests and the level of schooling at which they teach, i.e., elementary or secondary. For example, consistent with the finding reported earlier that high school teachers use tests included in curriculum materials less frequently than do elementary school teachers is the finding that elementary school teachers perceive commercial tests to be of higher quality than do secondary school teachers.

TABLE 1

USES TEACHERS MAKE OF STANDARDIZED TESTS

(1) In a recent large-scale study of test use in American schools (Burry et al., 1982), the majority of elementary school teachers and a sizable number of high school teachers reported that they regarded standardized tests as one of many sources of information important for planning instruction and grouping students at the beginning of the year, and for adjusting instructional groups during the year. Somewhat surprisingly, over 15 percent of the elementary school teachers reported they also used standardized test results to help determine students' report card grades. At the high school level, 12 percent of the English teachers and 8 percent of the math teachers surveyed also used standardized test results for grading.

(2) Teachers appear to make very limited use of standardized test results in setting the pace of instruction (Barr, 1975).

(3) It is unclear whether teachers view standardized tests as providing new and distinctive information, or whether such results are used primarily to confirm judgments that they have already made. Interviews with a sample of 68 elementary school teachers from both urban and suburban schools (Salmon-Cox, 1981) indicate that standardized test results are mainly used to confirm teacher judgments, and not to help develop them. However, the importance attached to standardized test results by teachers surveyed in the study by Burry and associates cited above is not entirely consistent with this conclusion.

(4) At least some teachers use results from standardized tests to correct negative misperceptions they may have developed about a student's ability. When test scores are lower than expected these teachers tend to discount the scores. When scores are higher than expected, this is an indication to the teachers that they have overlooked something (Salmon-Cox, 1981). Contrary to past claims, there is little evidence that test scores cause teachers to lower expectations for students.

(5) At the elementary school level, teachers appear to assign about the same importance to standardized test results in reading as in mathematics. At the high school level, English teachers appear to place considerably more importance on standardized tests than teachers of mathematics (Burry et al., 1982).

(6) Teachers whose students generally perform poorly on standardized tests are less likely to use test results than teachers whose students generally do well on such tests (Yeh, 1978).

(7) Teachers with more teaching experience are more likely to use standardized test results than are less experienced teachers (Yeh, 1978).

TABLE 2

TEACHER OBSERVATIONS

(1) The informal, minute-to-minute observations that teachers make in the classroom should be distinguished clearly from the planned observations of students' performance that characterize "performance tests." Teachers use both kinds of observations, but informal observations are much more likely to focus on student involvement in learning activities and the smoothness of classroom processes than on students' progress toward attaining desired learning goals (Clark & Peterson, 1981).

(2) One study indicates that teachers rely heavily on their own observations and students' classroom work to determine students' report card grades (Burry, et al., 1982). Another study indicates that information obtained through observation does not play a large role in teachers' grading (Stiggins & Bridgeford, 1982). The discrepancy between these findings may be based on the varied meanings teachers ascribe to the term "observation." Perhaps formal observation of student performance and formal evaluation of class exercises figure prominently in grading, whereas informal observations do not. This interpretation is consistent with evidence from the study by Clark and Peterson (1981) referred to above.

(3) With respect to both formal and informal observations, teachers channel information into at least three different categories: (1) information that calls for an immediate response, e.g., Sally begins to hit John on the head; (2) information that warrants delayed action and continued observation, e.g., Bob is "off task" for a few moments, but may resume work on his own shortly; and (3) information that needs to be stored for subsequent analysis, e.g., several students are having a hard time asking higher-order questions. Teachers often assign the same kind of information to different categories, e.g., one teacher may respond immediately to correct a student error, while another may continue to observe the student's response to see if an underlying error pattern emerges (Joyce, 1981).

(4) Though teachers rely a great deal on classroom observation to "fine tune" instruction, i.e., increase the time allotted to a particular activity, shift into a new activity, ask more questions about a topic or provide more examples of a concept, they rarely use assessment information as a basis for changing their general approach to teaching, e.g., using recitation-oriented vs. inquiry-oriented styles or a whole-group vs. small group management structure (Joyce, 1981).

(5) Teachers use a range of procedures for evaluating information obtained through planned observation, the most common of which seems to be checklists (Stiggins & Bridgeford, 1982).

(6) Teachers rarely ask another teacher to observe student performance or rate a student product, but peer and student self-ratings are widely used (Stiggins & Bridgeford, 1982).

(7) Considerable attention has been paid to the role that observation plays in teachers' assessment and interpretation of students' reading performance. It is evidently common for teachers who have not received special training to observe inappropriate and irrelevant aspects of children's reading behaviors as they attempt to diagnose the causes of students' reading difficulties (Polin, 1981).

(8) There is some evidence that teachers' observations of students' interpersonal classroom behavior are less accurate and stable than their perceptions of students' overall achievement level in a subject area (Elmore & Beggs, 1972).

TABLE 3

TEACHER ATTITUDES TOWARD AND CONCERNS ABOUT TESTING

(1) Teachers generally feel that testing motivates their students to study harder (Burry et al., 1982).

(2) Teachers by and large have a more positive attitude toward standardized tests than measurement specialists attribute to them (Stetz & Beck, 1982).

(3) Teachers generally feel that the direction of their districts' testing program are, to some extent, influenced by parental concerns about test results (Yeh, 1978).

(4) The more knowledge teachers have about testing, the more likely they are to regard tests as useful (Yeh, 1978).

(5) Teachers commonly assume that others outside the classroom, specifically parents, attach greater significance to standardized test results than they do (Salmon-Cox, 1981).

(6) Teachers generally feel that there is a close match between what they are teaching and the content of tests required by the state or district (Burry et al., 1982). Nonethless, teachers often express interest in "diagnostic" tests that not only are closely matched to the curriculum with which they are working but available to them on a day-by-day basis (Salmon-Cox, 1981).

(7) Teachers generally feel that tests developed in their district are very good (Burry et al., 1982).

(8) Teachers generally feel that tests of minimum competency/proficiency/functional literacy should be required of all students for promotion to certain grade levels, or for high school graduation (Burry et al., 1982).

(9) Elementary school teachers are more likely to perceive tests of minimum competency as unfair to particular students than are secondary school teachers (Burry et al., 1982).

(10) Elementary school teachers generally perceive commercial tests to be of higher quality then do secondary school teachers (Burry et al., 1982).

3. What do teachers generally know about testing, and what aspects of testing, or test usage, would they like to know more about?

Teachers knowledge about measurement is limited. Measurement rarely is a significant component of teacher preparation programs. When inservice training in testing is provided by a district it typically focuses on preparing for and administering district-required tests; little attention is given to the design or selection of classroom tests, or the use of test information. Secondary school teachers have received even less training in testing than elementary school teachers. Fortunately, there is evidence that teachers can and do benefit from inservice programs linking testing and instruction, at least programs that are built around realistic classroom situations and that include ample opportunity for teachers to exchange ideas with their colleagues. A summary of research on teacher's knowledge about and training in testing, and their expressed need for additional training, is presented in Table 4.

4. What advances in the technology of testing, if any, are likely to strengthen the link between testing and teaching?

Unfortunately, few of the recent technological developments in testing have had, or are likely to have in the near future, any impact on the testing practices of classroom teachers. There is little reason to believe that teachers will use the technical procedures for test item specification and production that psychometricians have developed. Nor do complex approaches to test review and validation stand a chance of finding their way into the professional repertoire of the typical teacher. Some of the efforts to build tests and test item pools that relate closely to what teachers are teaching may extend the resources for testing available to teachers. But whether these efforts will affect teachers' testing practices to any significant degree is an open question.

Table 5 summarizes key developments in the technology of testing. It also provides an overall assessment of the implications of these developments for teachers and teaching.

TABLE 4

TEACHERS' KNOWLEDGE ABOUT AND TRAINING IN ASSESSMENT, AND THEIR EXPRESSED NEEDS FOR ADDITIONAL TRAINING

(1) Teachers know little about basic measurement concepts (Rudman et al., 1980).

(2) Based on interviews with a small sample of teachers it would appear that most teachers either have never had a course in measurement or have had only one course. Fewer high school teachers have taken courses in measurement than teachers in middle or elementary schools (Stiggins & Bridgeford, 1982).

(3) The majority of elementary school teachers surveyed in a large-scale study (Burry et al., 1982) indicated that their district had provided inservice training in a number of areas related to testing, including procedures for administering tests required by the state or district, and for interpreting information coming from these tests; procedures other than tests to assess student achievement; and procedures for relating instruction more closely to the content covered on required tests. However, relatively few teachers reported receiving training in test construction or selection, or in using test results for instructional improvement.

(4) High school teachers evidently have received considerably less training or assistance from their districts in test-related matters than elementary school teachers. The one area in which a large number of high school teachers have received assistance is the administration and interpretation of "mandated" tests (Burry et al., 1982).

(5) Without specific training few teachers design evaluation procedures that are tied closely to instructional goals (Popham & Baker, 1970). Nor do they show particular skill in identifying the causes of students' low performance, at least in the area of reading (Polin, 1981). However, training in assessment and diagnosis does improve teachers' skills in these areas (Popham & Baker, 1970; Polin, 1981).

(6) A particularly promising approach to helping teachers use test information has been developed by Wanous and Mehrens (1981). This approach engages teachers in an analysis of a series of hypothetical, but realistic classroom situations that call for an instructional decision. Teachers are provided with a wide range of assessment information that is relevant to the decision under consideration. Exemplary ways of using this information to guide the decision are offered, against which teachers compare their own decision-making approach. Volunteer teachers received training using this approach in their own schools. The training provided extensive opportunity for teachers to work with colleagues on the problems posed in the hypothetical cases. By and large, teachers reported that the training helped them to design their own assessment procedures; to select more carefully curriculum packages that included assessment measures; and to use data more systematically in making instructional decisions (Wanous & Mehrens, 1981).

(7) Teachers at the elementary and junior high levels have expressed a need for more knowledge about measuring the affective and social dimensions of their classrooms. These teachers also have indicated an interest in extending their knowledge of assessment procedures in areas other than the familiar ones of mathematics and reading (Wanous & Mehrens, 1981).

TABLE 5

ADVANCES IN THE TECHNOLOGY OF TESTING AND THEIR IMPLICATIONS FOR TEACHERS AND TEACHING

(1) New procedures have been developed for designing tests that correspond closely to the scope and emphasis of instruction, as reflected in the instructional materials teachers are using (Hanson & Colleagues, 1980). These procedures have been used effectively by the Southwest Regional Educational Laboratory to develop program evaluation systems for local school districts. Information on students' learning in reference to a district's instructional scope and sequence generally is collected in September and at mid year, and used by administrators and teachers to identify priorities for instructional improvement.

(2) Steps have been taken to develop assessment procedures for teachers in areas that traditionally have not been assessed systematically, e.g., the resources for assessing speaking skills made available through the Speech Communication Association (Annadale, Virginia); the procedures for evaluating student writing samples described in publications of the Northwest Regional Educational Laboratory (Spandel & Stiggins, 1980) and the National Council of Teachers of English (Najmy, 1980); and tests like the "deductive logic and assumption recognition" tests prepared by the Instructional Objectives Exchange (Los Angeles). These assessment resources are intended for teachers' use. But the extent to which teachers know about or use these resources has yet to be determined.

(3) Procedures also have been developed for creating test item pools that are linked explicitly to the curriculum being implemented in a district (Valley Education Consortium, 1982). These pools permit teachers to generate tests based on what they are teaching at a level of specificity appropriate for a particular purpose, e.g., pretesting, progress-checking, posttesting, end-of-year testing. The pools also can be used by administrators to form tests suitable for program monitoring and evaluation. Evidence is not yet available, however, on the impact of curriculum-based test item pools on teachers' testing-teaching practices.

(4) Procedures for assuring quality in the construction of criterion-referenced tests and for analyzing test results have been refined and extended in recent years (Berk, 1980; Hambelton, 1980; Nitke, 1974). However, these procedures, such as those involved in reliability and construct validity investigations, or in using standard formulae for determining appropriate test lengths, do not seem to be used by teachers at present. There is no reason to believe that this situation will change in the near future, unless time and resources available to teachers for test design, scoring, and interpretation are dramatically increased.

(5) New procedures have been developed for defining the domain of outcomes a test is to measure. These include "amplified objectives," "item forms," and "test specifications." However, training programs have not been designed to teach teachers to use these procedures. The chief developer of practitioner-oriented domain-specification strategies, James Popham, reports that "he is unable to reduce the process to a form that is directly teachable (Popham, 1980, p. 17)." Even if teachers were trained in domain-specification strategies, it is not clear that they would have time to use them on a regular basis in their actual work.

(6) A number of highly technical procedures for generating test items from a domain have been developed, e.g., algorithms and linguistic transformations (Millman, 1980; Roid & Haladyna, 1982). These procedures, which commonly are used with the aid of computer technology, appear to have limited, if any, utility for the classroom teacher.

(7) Some progress has been made in assessing the underlying cognitive strategies students use to carry out various learning tasks in specific subject areas (Adams & Collins, 1979; Bereiter, 1979). Procedures for assessing cognitive strategies, and for identifying inappropriate or suboptimal uses of a strategy, are in an early stage of development. Preliminary procedures of this kind, however, have been incorporated with apparent success in several diagnostically-oriented instructional programs, most notably in Project Torque (Schwartz & Taylor, 1979), which focuses on the teaching of computation and measurement skills.

5. What specific approaches to integrating testing and teaching have been developed, and how applicable are they to regular high school classrooms?

A variety of specific approaches have been developed for integrating testing and instruction, most notably mastery learning and computer assisted instruction. Less commonly used approaches include adaptive testing, answer-until-correct testing methods, self-scoring and self-diagnosis procedures, and tests that simulate real-life problem situations. Brief descriptions of each of these approaches are presented in Table 6.

Specialized approaches to integrating testing and teaching probably will not be used in typical classrooms without the assistance of computer technology. These approaches call for a high level of individualized instruction. Under present conditions, particularly in high schools, this level of individualization cannot be managed effectively. Even mastery learning, which was developed independently of advances in micro-computer technology, poses serious classroom management problems for most teachers, for it adds a great deal of paperwork to their teaching load. Teachers will need computer assistance to manage the large quantity and refined quality of information on individual student achievement that most models for integrating testing and teaching call for.

Conclusion

A good deal of research has been done over the last decade that relates to the issue of integrating testing and teaching. For the most part this research is encouraging. It suggests that teachers find value in many types of tests, and use test information for a variety of purposes. It also suggests that teachers are aware of the limitations of tests, and use them selectively in combination with other sources of information on students. In addition, research indicates that teachers can increase their effectiveness in designing and using tests through inservice training programs.

Research and development efforts also suggest a note of caution. Scientific progress made in test design, scoring, and validation does not automatically translate into

TABLE 6

SPECIFIC APPROACHES TO INTEGRATING TESTING AND TEACHING

Adaptive Testing. This is an approach to testing in which the particular items administered to an examinee at one point in time depend upon the examinee's performance on items administered at a previous time. The number of items administered may vary, e.g., examinees who are clearly above or below some standard may be assessed with relatively few items. The difficulty level of the items may vary, e.g., examinees who are doing very well on one set of items are administered harder items. Computers often are used to facilitate adaptive testing (Millman, 1980).

Answer-Until-Correct (AUC) Testing Methods. These are testing procedures in which the student is permitted to continue selecting from item alternatives until the correct alternative is chosen. AUC procedures enable the teacher to review a student's entire response pattern and thereby gain a better understanding of the causes of low performance (Wilcox, 1982).

Computer-Assisted Instruction (CAI). In most CAI programs, testing is included in the form of pretests, progress tests, and mastery tests. The computer determines what each student should do next on the basis of test results; e.g., go on to more advanced material, receive additional instruction or practice, etc. (If remedial activities prescribed by the computer do not result in satisfactory performance, typically the computer tells the student to see the teacher). Some CAI programs match learning activities to the ability levels, learning style, e.g., auditory or visual, and cognitive style, e.g., abstract or concrete, of the learner (Charp, 1979). At present, CAI programs probably are used most extensively to teach the basic skills of mathematics, but CAI programs currently are being designed to develop more sophisticated skills (Martin, 1981).

Self-Scoring and Self-Diagnosis Procedures. These are procedures in which students score tests themselves, and are provided with an interpretive booklet that offers explanations of why each keyed answer is considered preferable to the alternatives. The College Board's Career Skills Assessment Program, for example, features tests of this kind. Answer sheets accompanying the program are designed so that as soon as the student has completed the test, the layered answer sheets can be separated. Students can immediately score their own answers on one part; the other part can be machine scored. The accompanying test booklet gives a rationale for each answer and presents illustrations of how the skills assessed on the test can be applied in practice.

Tests that Teach. Some tests are so closely connected to learning that they represent a form of teaching. The tests presented to pilot trainees by flight simulators, for example, are excellent examples of tests that teach. The trainees learn how to face emergencies, respond quickly and appropriately, and see the consequences of their actions immediately (Krumboltz, 1982). Simulators for other learning problems are being developed, e.g., the Career Decision Simulation constructed by Krumboltz, Hamel, and Scherba (Krumboltz, 1982).

Mastery Learning. In mastery learning the curriculum is clearly sequenced so that foundation concepts and skills are well developed before advanced content is introduced. Students are assessed prior to instruction to determine precisely at what point in the curriculum they should begin work. Instructional time is free to vary to accommodate the different rates at which students learn. Similarly, a variety of materials and strategies is used to accommodate differences in student learning styles. Tests are used routinely to monitor student progress toward attaining the learning outcomes identified in the curriculum, and to identify needs for instructional adjustment. Mastery tests are given upon completion of an instructional unit. These tests indicate whether a student is ready for new work, or whether instructional "recycling" is needed. Tests thus play a key role in mastery learning in deciding what students should learn and when. The effectiveness of mastery learning has been demonstrated in a number of studies (Block, 1974; Bloom, 1976), including a study of a mastery learning approach to teaching high school chemistry (Swanson & Denton, 1976).

benefits for the classroom teacher. Nor do technically sophisticated approaches to integrating testing and teaching have direct utility for the typical teacher. Research-based procedures for strengthening the link between teaching and testing will have to be carefully adapted for use in the nation's schools and classrooms.

SJ129

References

Adams, M. & Collins, A. A schema-theoretic view of reading. In R. Freedle (ed.), New directions in discourse processing. Norwood, NJ: Ablex, 1979.

Barr, R. How children are taught to read: Grouping and pacing. School Review, 1975, 83, 479-498.

Bereiter, C. Development in writing. In L. Gregg and E. Steinberg (Eds.), Cognitive processes in writing. Hillsdale, NJ: Lawrence Erlbaum Associates, 1979.

Berk, R. Item analysis. In R. Berk (Ed.) Criterion-referenced measurement. Baltimore, MD: The Johns Hopkins University Press, 1980.

Block, J. Schools, society, and mastery learning. New York: Holt, Rinehart, and Winston, 1974.

Bloom, B. Human characteristics and school learning. New York: McGraw Hill, 1976.

Burry, J., Catterall, J., Choppin, B., & Dorr-Bremme, D. Testing in the nation's schools and districts: How much? What kinds? To what ends? At what costs? Report No. 194 Los Angeles, CA: Center for the Study of Evaluation, University of California, Los Angeles, 1982.

Charp, S. Case study: Instruction and testing, school district of Philadelphia. In R. Tyler and S. White (Chair), Testing, teaching and learning: Report of a conference on research on testing. Washington, D.C.: U.S. Department of Health, Education and Welfare, National Institute of Education, 1979.

Clark, C. & Peterson, P. Stimulated recall. In B. Joyce, C. Brown, & L. Peck (Eds.) Flexibility in teaching. New York: Longman, 1981.

Elmore, P. & Beggs, D. Stability of teacher ratings of pupil behavior in a classroom setting. Paper presented at a meeting of the American Personnel and Guidance Association. March 1972.

Hambleton, R. Test score validity and standard-setting methods. In R. Berk (Ed.), Criterion-referenced measurement. Baltimore, MD: The Johns Hopkins University Press, 1980.

Hanson, R., Behr, G., Meguro, B., & Bailey, J. Development and verification of instructionally sensitive achievement tests. Technical report 69. Los Alamitos, CA: Southwest Regional Laboratory, December, 1980.

Joyce, B. Toward a theory of information processing in teaching: The South Bay study. In B. Joyce, C. Brown, & L. Peck (Eds.), Flexibility in Teaching. New York: Longman, 1981.

Krumboltz, J. Tests and guidance: What students need. In W. Schrader (Ed.), Measurement, guidance and program improvement. San Francisco, CA: Jossey-Bass, 1981.

Martin, J. On reading, writing, and computers: A conversation with John Henry Martin. Educational Leadership. October, 1981, 60-64.

Millman, J. Computer-based item generation. In R. Berk (Ed.), Criterion-referenced measurement. Baltimore, MD: The Johns Hopkins University Press, 1980.

Najimy, N. (Ed.) Measure for measure: A guidebook for evaluating students' expository writing. Urbana, IL: NCTE, 1981.

Nitko, A. Problems in the development of criterion-referenced tests: The IPI Pittsburgh experience. In C. Harris, M. Alkin, & W. Popham (Eds.), Problems in criterion-referenced measurement. Los Angeles, Center for the study of Evaluation, University of California, 1974.

Polin, R. A study of preceptor training of classroom teachers in reading diagnosis. East Lansing, Institute for Research on Teaching, Michigan State University, 1982.

Popham, J. Domain specification strategies. In R. Berk (Ed.), Criterion-referenced measurement. Baltimore, MD: The Johns Hopkins University Press, 1980.

Resnick, L. Introduction: Research to inform a debate. Phi Delta Kappan, May, 1981, 623-624.

Roid, G. & Haladyna, T. A technology for test-item writing. New York: Academic Press, 1982.

Rudman, H., Kelly, T., Wanous, D., Mehrens, W., Clark, C., & Porter, A. Integrating assessment with instruction: A review (1922-1980). East Lansing, MI: Institute for Research on Teaching, Michigan State University, 1980.

Stetz, F. & Beck, M. Attitudes toward standardized tests: Students, teachers, and measurement specialists. Measurement in Education, 1981, 12, 1-10.

Stiggins, R. & Bridgeford, N. Final research report on the role, nature, and quality of classroom performance assessment. Portland, OR: Northwest Regional Educational Laboratory, November, 1982.

Swanson, P. & Denton, J. A comparison of remediation systems affecting achievement and retention in mastery learning, 1976 (ERIC Document Reproduction Service No. Ed 131 037).

Valley Education Consortium, Program guides for basic skill development in mathematics, reading and writing. Monmouth, OR: Valley Education Consortium, Western Oregon State College, 1982.

Wanous D. & Mehrens, W. Helping teachers use information: The data box approach. Measurement in education, 12, 1981, 1-10.

Wilcox, R. Some new results on an answer-until-correct scoring procedure. Journal of Educational Measurement, 1982, 19, 67-74.

Yeh, J. Test use in schools. Washington, D.C.: U.S. Department of Health, Education and Welfare, National Institute of Education, 1978.

APPENDIX 3

A Questionnaire Used to Assess High School Teachers' Testing Practices

A Questionnaire Used to Assess High School Teachers' Testing Practices

The questionnaire presented on the pages that follow was designed to help us gain a better understanding of high school teachers' approach to testing and the use of test information. It also was intended to yield information on the test-related resources available to teachers and the attitudes and concerns teachers have about testing.

The questionnaire was completed in the spring of 1983 by 526 teachers in 9 high schools in Oregon and 6 schools outside the state. The schools in the sample consisted of some of the schools in which members of the Teacher Advisory Panel taught and some which were involved in a goal-based education project sponsored by the Northwest Regional educational laboratory. All full-time teachers in the sample schools were asked to complete the survey. Approximately 64 percent did so.

Results from the survey were tabulated in May of 1983. The principal of each school in the sample received a summary of results for his or her building that same month. A summary of responses to each item on the questionnaire is available upon request from the Teaching Research Division of the Oregon State System of Higher Education.

Unfortunately, the information produced by the survey provided only limited help in developing the Handbook. This was largely because the results simply confirmed that teachers use a variety of test item formats, place major emphasis on the use of tests for grading, generally have limited formal training in assessment, and so forth. In restrospect, it might have been better to give the questionnaire to teachers who are known to integrate instruction and assessment in a thorough way and those who are known to make only weak connections between instruction and assessment. This may have helped to enhance our understanding of the practices and attitudes associated with different approaches to teaching and testing.

The survey did, however, provide a basis for developing the teacher survey instrument that has been used in the CEPM research project described earlier in this report. More information about the refined version of the teacher survey and its use will be provided in subsequent reports to the NIE, made in connection with the CEPM project.

TEACHING RESEARCH DIVISION
OREGON STATE SYSTEM OF HIGHER EDUCATION

A Survey of High School Teachers' Testing Practices and Attitudes Toward Testing

Background

This survey asks you about your approach to student testing. It also asks about your use of test results. For purposes of the survey, testing is considered as any systematic way of gathering information on students.

Findings from the survey will be included in a handbook for high school teachers on linking testing and teaching. The handbook is being developed by Teaching Research, a state-supported agency specializing in educational research and development, through contract with the National Institute of Education. It is intended that the handbook will help high school teachers make better use of tests as aids in instructional planning and decision-making. A copy of the handbook will be sent to your principal's office when it is completed, in the spring of 1984. The handbook will contain a summary of results found through the survey. A preliminary report of results will be sent to you toward the end of this school year.

Directions

Please answer the questions candidly. Information is needed on what you do - not on what you feel you should do. It is recognized that no teacher can do all that he or she would like to do. Also, do not put your name on the survey; responses are to be anonymous.

Thanks for your willingness to complete the survey.

Testing Formats and Procedures

1. Which of the following tests and assessment procedures do you generally use? (Check all that apply.)

_____ final exams	_____ in-class exercises
_____ mid-terms	_____ projects and reports
_____ unit tests	_____ homework assignments
_____ quizzes	_____ other (please specify) _____

_____

2. Please indicate the frequency with which you use each of the following test formats (check the appropriate space).

Format	Frequency of Use			
	Very Often	Often	Occasionally	Seldom
• multiple choice	—	—	—	—
• true/false	—	—	—	—
• matching	—	—	—	—
• fill in the blank or problem-solving	—	—	—	—
• short-answer essay	—	—	—	—
• extended essay	—	—	—	—
• other student products, e.g., a painting, musical composition, mechanical drawing	—	—	—	—
• student performance, e.g., answering questions out loud; typing a letter; dribbling a basketball	—	—	—	—
• open book tests, i.e., tests in which students can use their notes or books	—	—	—	—
• take home tests, i.e., tests that students do at home	—	—	—	—

3. Please indicate the frequency with which you use additional sources of information about student learning or student characteristics.

Source of Information	Frequency of Use: Very Often	Often	Occasionally	Seldom
• previous teachers' records, reports or grades	___	___	___	___
• information provided by a student's parents	___	___	___	___
• standardized tests of achievement	___	___	___	___
• tests of student aptitudes	___	___	___	___
• state or district developed tests	___	___	___	___
• measures of student attitudes or interests	___	___	___	___
• measures of student learning styles	___	___	___	___
• other (please specify)	___	___	___	___

4. Which of the following practices do you use on a regular basis to help students prepare for your tests? (Check all that apply.)

_____ provide a review of the material to be covered on the test

_____ give a practice test

_____ construct the test using questions that have already been discussed in class

_____ provide training in test-taking skills

_____ other (please specify) _____

5. Which of the following testing practices do you use on a regular basis when working with low-ability students? (Check all that apply.)

_____ tell the students not to worry if they don't do very well on a test

_____ give the students special tests matched to their ability level

_____ give the students extra help so they have a better chance of doing well on tests

_____ give the students less advanced work and correspondingly easier tests

_____ other (please specify) _____

6. How much class time do your students spend in a typical week taking tests and reviewing test results? (Check the appropriate space.)

_____ 0 - 30 minutes

_____ 30 - 60 minutes

_____ 60 - 90 minutes

_____ over 90 minutes

Analyzing, Managing and Reporting Test Information

7. When analyzing results from objective tests, I generally: (check the appropriate space)

_____ determine each student's overall score on the test, but do not perform any kind of item analysis; i.e., I rarely look at how students do on individual test items

_____ analyze responses to a few items in which I am particularly interested

_____ analyze responses to each test item

_____ I don't give objective tests

8. Which of the following resources for scoring and managing test results, if any, are available to you: (check all that apply)

_____ paid adult or student aides

_____ volunteer adults or students

_____ scoring machine

_____ in my classroom

_____ in the building

_____ can use for any objective test I give

_____ can use only for some exams (e.g., final)

_____ in another building

_____ can use for any objective test I give

_____ can use only for some exams (e.g., final)

_____ other (please explain) _____

9. How much time do you typically spend each week outside of class scoring tests or grading papers or projects? (Please estimate to nearest 15 minutes.)

_____

10. Listed below are three types of measures of student learning. About how long after students complete each of these measures do you generally let them know how they did on it? (Check the appropriate spaces. If you do not regularly use one of the measures listed, leave the spaces next to it blank.)

	Within 24 Hours	Within 2-3 Days	Within a Week	Within 2 Weeks
Quizzes	___	___	___	___
Unit Tests	___	___	___	___
Mid-Terms	___	___	___	___

11. When discussing a student's academic progress with his/her parents, which of the following practices, if any, do you usually follow? (Check all that apply.)

_____ tell the parents the student's scores on various tests that you have given

_____ show the parents examples of the answers that the student has written to your test questions

_____ indicate how the student's scores on your tests compare to the class average

_____ tell parents how the student has performed on general aptitude

Using Test Information

12. In what ways do you use test information? (Check all that apply.)

_____ to group students for instruction, or particular learning activities

_____ to decide whether to teach the class more about a topic or to go on to another topic

_____ to decide whether a particular student knows enough that he or she can skip a unit

_____ to decide whether a particular student needs extra help

_____ to decide what type of teaching approach to implement, e.g., lecture vs. discussion; small group vs. whole class activity, etc.

_____ to assign grades to students

_____ to evaluate the effectiveness of a course

_____ other (please specify) _____

_____

_____

13. Which of the following practices do you use on a regular basis when a large percentage of your class does poorly on a test? (Check all that apply.)

_____ reteach the class the material on which the test was based

_____ reteach only those students who did poorly

_____ allow students to retake the test, or a comparable test

_____ ask students who did well on the test to provide some form of in-class help or tutoring to those who did poorly

_____ move on to the next unit anyway

_____ other (please specify) _____

_____

_____

14. Listed below are a number of factors that might explain why a large percentage of a class does poorly on a test you have prepared. Place a check on the line that indicates your perception of the frequency with which each of these factors comes into play.

	Frequently a cause of poor performance	Sometimes a cause of poor performance	Seldom a cause of poor performance
TIME: Not enough time was available to teach the concepts and skills thoroughly, or to provide students with sufficient opportunity to practice or apply them.	_____	_____	_____
FEEDBACK PROCEDURES: Too little information was provided to students on their progress toward attaining the learning outcomes that were covered on the test.	_____	_____	_____
METHODS/MATERIALS: The instructional methods and materials used were not well suited to the students, or to the type of learning outcomes expected.	_____	_____	_____
CONTENT: The concepts and skills that students were expected to learn were too complex, too abstract, or too many in number.	_____	_____	_____
TEST DESIGN: The test did not match closely what was taught, or required students to demonstrate what they had learned in unfamiliar ways.	_____	_____	_____

15. Which of the following practices do you use on a regular basis when a student consistently does poorly on your tests? (Check all that apply.)

_____ talk with the student

_____ talk with other teachers about the student

_____ make a point to watch the student during class periods

_____ give the student extra help

_____ give the student a diagnostic test to see if he or she has the prerequisite skills needed to do the regular coursework

_____ give the student less advanced work

_____ talk to the student's counselor or parents

_____ explore the possibility of the student's transferring to a less advanced course

16. Listed below are a number of factors that might explain why an INDIVIDUAL STUDENT consistently does poorly on tests you have prepared. Place a check on the line that indicates your perception of the frequency with which each of these factors comes into play.

	Frequently a cause of poor performance	Sometimes a cause of poor performance	Seldom a cause of poor performance
The student was frequently absent or tardy during the time the concepts or skills covered on the test were taught.	____	____	____
The student did not participate fully or consistently in learning activities, or was distracted from learning by disruptions in the classroom environment.	____	____	____
The student did not complete homework assignments.	____	____	____
The student lacks the prerequisite concepts or skills needed to succeed in the course.	____	____	____
The student has difficulty learning concepts or skills through the type of instructional method or materials used with the class as a whole, or within the time constraints normally imposed.	____	____	____
The student has no interest in the topics being taught.	____	____	____
The student has difficulty demonstrating what he/she knows or can do in the form called for on the test, quiz, or exercise.	____	____	____
The student was ill or unduly distracted the day the test was given.	____	____	____
The student needed more time to finish the test.	____	____	____

17. Which of the following practices do you use on a regular basis when a student consistently scores at the top of the class on your tests? (Check all that apply.)

_____ talk with the student

_____ talk with other teachers about the student

_____ give the student more advanced work

_____ ask the student to tutor other students in the class

_____ give the student more opportunity to work independently

_____ talk to the student's counselor or parents

_____ explore the possibility of the student's transferring to a more advanced class

_____ other (please specify) _____

_____

_____

18. Do you use tests that accompany textbooks on a regular basis?

_____ Yes _____ No

If yes, which of the following uses do you generally make of these tests? (Check all that apply.)

_____ to supplement information from your own tests

_____ to use instead of your own tests

_____ to provide guidelines for designing your own tests

_____ other (please specify) _____

_____

_____

19. Listed below are a number of factors that teachers consider when assigning grades. Enter in front of each factor the weight it carries in most of your courses. Make the percentages add up to 100. If you do not use a source that is listed, leave the space in front of it blank.

_____ % final exam

_____ % mid-term

_____ % unit tests

_____ % projects or reports

_____ % homework assignments

_____ % quizzes or in-class exercises

_____ % performance on a standardized test

_____ % performance on a district-administered test

_____ % class participation

_____ % effort made by the student to learn

_____ % other (please specify) _____

_____

_____

20. Do your testing practices, including your use of test information, differ from one course to another?

_____ Yes _____ No

If yes, please explain what you do differently in each course, and why.

Concerns About Testing

21. Some teachers have concerns about the use of tests in their classrooms. Others don't. Check the concerns listed below that you have at this time.

_____ Students often seem uninterested in learning anything that they are not going to be tested on or assigned a grade for.

_____ The scoring and analysis of tests is quite time consuming.

_____ Students who do poorly on tests often feel they are stupid and have little incentive for studying hard.

_____ Many valuable aspects of learning do not seem to be measurable through tests, e.g., creativity, wisdom, fairness.

_____ Students put too much emphasis on how their test scores compare with others in the class.

_____ Too much instructional time is consumed by tests required by the state or federal government, or by the district.

_____ other (please specify) _____

_____

_____

22. Some teachers have concerns about the use of tests in education generally. Others don't. Check the concerns listed below that you have at this time.

_____ The public puts too much emphasis on national tests, such as the SAT, that aren't good measures of what is being taught in individual courses.

_____ Tests of student achievement are sometimes viewed as a means for monitoring teacher performance without taking into account the different ability levels and backgrounds of the students teachers have to deal with.

_____ The money spent on testing would be put to better use if it were used to purchase instructional materials.

_____ other (please specify) _____

_____

_____

Interest in Learning More About Test-Related Issues and Practices

23. How much interest do you have in attending a one-day workshop on designing and using tests as an aid to instructional planning and decision-making?

No Interest	Some Interest	Considerable Interest	Great Interest
_____	_____	_____	_____

24. Regardless of how you answer the previous question, please rank the following possible workshop topics in order of interest to you. Place a 1 in front of the topic of most interest, a 2 in front of the topic of next interest, and so on.

_____ how to design tests that are matched closely to what I have taught

_____ how to design tests in areas of learning that are generally considered to be hard to measure, e.g., creativity, wisdom, leadership

_____ how to use test results as a guide for planning remedial or enrichment activities

_____ how to score and manage test information with the aid of a micro-computer

_____ how to report test results to students, parents, or administrators

_____ other (please specify) _____

_____

Background Information

25. What subject do you teach? (Check the appropriate space.)

_____ Career Education	_____ Math
_____ English	_____ Physical Education
_____ Fine Arts	_____ Science
_____ Foreign Language	_____ Social Studies
_____ Health	_____ Vocational Education
_____ Home Economics	_____ Other (please specify) _____

26. How many years have you been teaching? _____

27. What professional degree do you hold?

_____ BA or BS

_____ BA or BS + 30 hours

_____ Master's Degree

_____ Master's + 30 hours

_____ Ph.D. or Ed.D.

28. How many courses in test design and measurement have you taken in preparing to be a teacher?

_____

When were these courses taken? _____

29. Have you participated in any workshops or other inservice training in testing or use of test information?

_____ Yes _____ No

If yes, please describe the workshops or training sessions and indicate when you participated in them.

_____

_____

_____

30. Please write the names of your school and district.

School: _____ District: _____

For Further Information
about the program and
related materials
write

THE OREGON EDUCATIONAL
COOPERATIVE PRESS
Todd Hall
Western Oregon State College
Monmouth, Oregon
97361

or call

Area Code 503-838-1220
Ext. 391, 392 or 393

The **Handbook For Integrating Teaching and Testing in the High School**, and related staff development materials, were developed by Drs. Glen Fielding and Del Schalock of the Teaching Research Division, Oregon State System of Higher Education. These resources for staff development build on a long history of work in Oregon on the integration of teaching and testing, and the use of test information in program management at all levels of schooling. It also builds on a long history of work in Oregon on school-based staff development, and on the work of the Valley Education Consortium in both of these areas. Both the **Handbook** and the **Staff Development Guides** were pilot tested in Oregon and Washington schools.

INTEGRATING TEACHING AND TESTING

A Staff Development Program

For Improving the Quality of Instruction

In High Schools

Developed under contract with

The National Institute of Education

1984

A STAFF DEVELOPMENT PROGRAM FOR TEACHERS AND ADMINISTRATORS

Created to help local districts take advantage of recent advances in instruction, testing and the use of test information

- in aiding student learning;
- in guiding instruction;
- in developing a curriculum and preparing plans for a course;
- in evaluating and improving instructional programs.

Provides the understanding and skills needed for departments to effectively implement instructional programs that depend heavily on testing and the use of test information, for example

- mastery learning;
- goal-based, or outcome-based instruction;
- individualized instruction.

Emphasizes the emerging concept of curriculum alignment in testing, and builds on the established technology of criterion referenced testing.

A LOW-COST PROGRAM THAT CAN BE IMPLEMENTED WITH EXISTING PERSONNEL

Designed for use by a faculty of a department, but may be used by more than one department at a time.

Operated by the school's Principal, a Department Chair, a "Lead Teacher," and a resource person in or outside the district who is knowledgeable about recent developments in testing.

The basic training program requires only two days of released time for teachers. Three additional days can be devoted to advanced training, but these are optional.

The basic training program gives teachers the understanding and skills they need to start where students are, monitor learning gains that are made, and plan or adapt instruction accordingly.

The basic program takes as its focus one or more "units" of instruction currently being taught. Instructional units are prepared during the training program that fully integrate teaching, testing and the use of test information.

The advanced training program focuses on an entire course, and on the connection between courses and programs of instruction.

A COMPLETE SET OF PROGRAM RELATED RESOURCES

The staff development program makes use of

- a **Handbook** for integrating teaching and testing in high school programs;
- a **Resource Guide for Teachers** that is designed to be used by members of a department as a team; and
- a **Program Planning Guide for Administrators and Lead Teachers** that is designed for Superintendents and other central office administrators responsible for **initiating** the program, and Principals, Lead Teachers and Department Chairs who are responsible for **implementing** the program.

In combination these documents make up an integrated, self-contained training system that can be used by local districts to accommodate local needs and resources.

4/15/85

An Outline of the Evolution and Completion of Contract #400-82-0013: Developing a Teachers' Handbook for Integrating Teaching and Testing in High School

Products Called for in Original Contract	Modifications Made and Approved During the Course of the Project	Date Submitted to the NIE
1. Literature Review	None	. January, 1983 . Additional copies were included in the Final Report, submitted February, 1985.
2. Survey of high school teachers' testing practices	Originally it was thought that information from the survey would be used as a guide in designing the Handbook. As it turned out, the survey results had limited utility in preparing the Handbook, but the survey has proved very useful in a related research study sponsored by the NIE. (See notes in the right-hand column.)	. The final version of the survey was submitted in Spring, 1983. . Information on the use and refinement of the survey is contained in the Final Report. Related information has been provided in progress reports submitted to the NIE in connection with a project on high school testing sponsored by CEPM through Grant NIE-G-81-0110.
3. Teachers' Handbook	The Handbook went through a more extensive process of review and revision then originally planned (see pages 4-10 of the Final Report).	The final version of the Handbook was submitted in February, 1985.
4. Exemplary teaching-testing materials	Originally we were to work with teachers in developing a separate set of illustrative teaching-testing procedures. We did work with teachers in designing such procedures, but rather than treating them as separate documents we have integrated them in the Handbook and in the Resource Guide for Teachers. The Handbook and Resource Guide contain numerous examples of different types of learning goals, goal-based instructional activities and assessment procedures, and ways of using test information in the instructional process.	Included in the Teachers' Handbook and the Resource Guide for Teachers. (The final version of the Resource Guide accompanies this outline and is described below.)

Products Called for in Original Contract	Modifications Made and Approved During the Course of the Project	Date Submitted to the NIE
5. A description of how the Teachers' Handbook could be implemented in school-based inservice programs.	It was clear from the outset of the project that we needed to provide guidelines on using the Teachers' Handbook in school-based staff development programs. However, it was not until the second phase of the project that we recognized the need for a comprehensive, step-by-step implementation manual for teachers. While the Teachers' Handbook serves as a mini-text, the Resource Guide for Teachers that was developed through the no-cost extension of the project provides sequenced learning activities. The Resource Guide represents a more elaborate and thoroughgoing response to the issue of implementation than we originally anticipated. It also goes beyond a focus on individual classrooms and contains guidelines for integrating teaching and testing at the program-level.	The field test version of the Resource Guide for Teachers accompanies this outline. The Resource Guide will be further tested and refined during the remaining months of the CEPM/NIE grant NIE-G-81-0110, and submitted to the NIE in final form at the completion of that grant.
6. Administrator's Guide	The original contract called for the development of an administrator's guide, which would help district and building administrators plan and support local inservice training programs based on the Teachers' Handbook. As the project unfolded it became clear, however, that administrators would need to work hand-in-hand with "lead" or "master" teachers in structuring and carrying out the training, and in sustaining its effects. We pilot tested an inservice program in which lead teachers, with support from their principals, facilitated inservice workshops in their buildings. Based on the success of the pilot test, we decided to expand our original notion of an administrator's guide into a Planning Guide for Lead Teachers and Administrators. This Planning Guide does all that the original administrator's guide was to do, but also identifies major roles and responsibilities for lead teachers.	A first draft, "field review" version of the Planning Guide for Administrators and Lead Teachers accompanies this outline. It will be more fully developed and pilot tested during the remaining months of the CEPM/NIE grant NIE-G-81-0110, and submitted to the NIE in final form at the completion of that grant.
7. Final Report	_____	This was submitted in February, 1985.